HHS Public Access

Author manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2020 October 19.

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2020 February ; 29(2): 477–486.

doi:10.1158/1055-9965.EPI-19-0755.

Identification of Novel Loci and New Risk Variant in Known Loci for Colorectal Cancer Risk in East Asians

A full list of authors and affiliations appears at the end of the article.

Abstract

Background—Risk variants identified so far for colorectal cancer explain only a small proportion of familial risk of this cancer, particularly in Asians.

Methods—We performed a genome-wide association study (GWAS) of colorectal cancer in East Asians, including 23,572 colorectal cancer cases and 48,700 controls. To identify novel risk loci, we selected 60 promising risk variants for replication using data from 58,131 colorectal cancer cases and 67,347 controls of European descent. To identify additional risk variants in known colorectal cancer loci, we performed conditional analyses in East Asians.

Results—An indel variant, rs67052019 at 1p13.3, was found to be associated with colorectal cancer risk at $P = 3.9 \times 10^{-8}$ in Asians (*OR* per allele deletion = 1.13, 95% confidence interval = 1.08–1.18). This association was replicated in European descendants using a variant (rs2938616) in complete linkage disequilibrium with rs67052019 ($P = 7.7 \times 10^{-3}$). Of the remaining 59 variants, 12 showed an association at $P < 0.05$ in the European-ancestry study, including rs11108175 and rs9634162 at $P < 5 \times 10^{-8}$ and two variants with an association near the genome-wide significance level (rs60911071, $P = 5.8 \times 10^{-8}$; rs62558833, $P = 7.5 \times 10^{-8}$)in the combined

Corresponding Author: Wei Zheng, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, 2525 West End Avenue, 8th Floor, Nashville, TN 37203-1738. Phone: 615-936-0682; wei.zheng@vanderbilt.edu.

Authors' Contributions

Conception and design: Y. Lu, J. Kim, A. Shin, D.-H. Kim, Y.-O. Ahn, Z.-Z. Pan, A. Lindblom, U. Peters, W. Zheng

Development of methodology: Y. Lu, Z.-Z. Pan, W. Zheng

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S.-S. Kweon, C. Tanikawa, X.-O. Shu, W.-H. Jia, T.A. Harrison, J. Kim, A. Shin, D.-H. Kim, K. Matsuo, S.H. Jee, J. Shi, N. Wang, M.-H. Shin, H.-L. Li, Z. Ren, J.H. Oh, I. Oze, Y.-O. Ahn, J. Gao, Y.-T. Gao, Z.-Z. Pan, A.T. Chan, A. Gsur, J. Hampe, L.L. Marchand, L. Li, A. Lindblom, V. Moreno, P.A. Newcomb, K. Offit, P.D.P. Pharoah, F.J.B. van Duijnhoven, B.V. Guelpen, P. Vodicka, A. Wolk, A.H. Wu, U. Peters, K. Matsuda, W. Zheng

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Lu, X.-O. Shu, J.R. Huyghe, A. Shin, S.H. Jee, X. Guo, W. Wen, B. Li, Y.-O. Ahn, Z.-Z. Pan, Y. Kamatani, L. Hsu, J. Long, U. Peters, W. Zheng

Writing, review, and/or revision of the manuscript: Y. Lu, Q. Cai, X.-O. Shu, Y.-B. Xiang, J.Kim,A.Shin, D.-H.Kim,K. Matsuo, S.H. Jee, H.-L. Li,J.Gao, Y.-T.Gao, Z.-Z. Pan, A.T. Chan, A. Gsur, L.L. Marchand, L. Li, A. Lindblom, V. Moreno, P.A. Newcomb, P.D.P. Pharoah, F.J.B. van Duijnhoven, B.V. Guelpen, P. Vodicka, S.J. Weinstein, A. Wolk, A.H. Wu, L. Hsu, U. Peters, W. Zheng

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y.-B. Xiang, A. Shin, S.H. Jee, J. Shi, N. Wang, H.-L. Li, J.H. Oh, K.J. Jung, J. Gao, Y.-T. Gao, Z.-Z. Pan, S.J. Weinstein, Y.-X. Zeng, W. Zheng

Study supervision: J. Kim, Z.-Z. Pan, U. Peters, W. Zheng

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Disclosure of Potential Conflicts of Interest

Y. Kamatani reports receiving speakers bureau honoraria from Illumina Japan. No potential conflicts of interest were disclosed by the other authors.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

analyses of Asian- and European-ancestry data. In addition, using data from East Asians, we identified 13 new risk variants at 11 loci reported from previous GWAS.

Conclusions—In this large GWAS, we identified three novel risk loci and two highly suggestive loci for colorectal cancer risk and provided evidence for potential roles of multiple genes and pathways in the etiology of colorectal cancer. In addition, we showed that additional risk variants exist in many colorectal cancer risk loci identified previously.

Impact—Our study provides novel data to improve the understanding of the genetic basis for colorectal cancer risk.

Introduction

Colorectal cancer is the third most commonly diagnosed cancer in men and the second in women, with 1.65 million new cases and almost 835,000 deaths in 2015 (1). Inherited genetic susceptibility contributes significantly to the etiology of colorectal cancer (2). Rare high-penetrance germline mutations in colorectal cancer predisposition genes (*APC*, *MUTYH*, *MLH1*, *MSH2*, *MSH6*, *PMS2*, *PTEN*, *STK11*, *GREM1*, *BMPR1A*, *SMAD4*, *POLE*, *POLD1*, *NTHL1*, and *TP53*; refs. 3–5) are estimate to account for less than 10% of colorectal cancer cases in the general population (3). Over the past decade, genome-wide association studies (GWAS) have identified about 100 independent loci associated with colorectal cancer risk (3, 6–13). These common genetic risk variants, however, explain only a small proportion of the familial relative risk of colorectal cancers (10, 11, 13), indicating that additional susceptibility variants remain to be identified.

Asians differ significantly from European descendants in genetic architectures. Genetic studies in Asians may provide an opportunity to explore the genetic architecture of colorectal cancer including identification of novel variants. We established the Asia Colorectal Cancer Consortium (ACCC) in 2010 to identify new genetic risk factors for colorectal cancer. Over the past 10 years, we have identified about 30 novel colorectal cancer risk loci (6–8, 11, 14). To further increase the statistical power of uncovering novel susceptibility loci for colorectal cancer, we utilized data from studies of 58,131 cases and 67,347 controls of European ancestry (10) to replicate promising risk variants identified in GWAS of 23,572 cases and 48,700 controls recruited from 15 studies conducted in three East Asian Countries (China, Japan, and Korea). Furthermore, we performed conditional analyses to identify potential independent signals at each of the colorectal cancer risk loci identified in previous studies in Asians.

Materials and Methods

Overview of study population and study design

We recently reported results from a large GWAS conducted in the Asia Colorectal Cancer Consortium (ACCC) that identified 13 novel genetic loci for colorectal cancer risk (11). In this study, we increased the sample size further by including one additional study—the Korean-National Cancer Center CRC Study 2 (Korea-NCC2: 622 cases and 832 controls). Included in the current analysis were 23,572 colorectal cancer cases and 48,700 controls from 15 studies conducted in China, Japan, and South Korea (Supplementary Table S1;

Supplementary Data). To increase the statistical power, we used data from European descendants to replicate promising findings from the ACCC.

On the basis of the results from the meta-analysis of all Asian studies, we selected 60 promising variants ($P < 5 \times 10^{-3}$, Supplementary Table S2) that are 500 kb away from any established colorectal cancer risk loci (10, 11) at the time of study design for replication using data from 58,131 cases and 67,347 controls of European descent (10). These cases and controls were derived from three colorectal cancer study consortia: the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), the Colorectal Transdisciplinary (CORECT) Study, and the Colon Cancer Family Registry (CCFR). The details of each included study have been reported previously (9, 10). For this analysis, the Haplotype Reference Consortium panel without indel variants was used as reference for imputation.

All study protocols were approved by the relevant Institutional Review Boards (10, 11) and informed consents were obtained from study participants. Research was conducted in accordance with the Belmont report.

Genotyping and imputation

Details of genotyping, genotype calling, quality control and imputation for the ACCC have been reported previously (Supplementary Data; refs. 6–8, 11, 14). The genotypes for samples from six studies (Shanghai-4, Aichi-2, Korea-NCC, Korea-NCC2, Korea-Seoul, and HCES2-CRC; Supplementary Table S1) genotyped with Illumina MEGA-Expanded Array (Illumina Inc.) were updated on the basis of the manually reclustered genotype cluster files (for 123K SNPs). Little evidence of population stratification was found in these studies (6, 8, 11), based on principal component (PC) analysis, using EIGENSTRAT (Supplementary Fig. S1; ref. 15). To increase the genome coverage and facilitate the meta-analysis, the 1000 Genomes Project phase III mixed reference haplotypes (version 5) were used to impute untyped genotype data with Michigan Imputation Server (minimac3 for imputation and SHAPEIT for prephasing).

We evaluated the quality of imputation using whole-genome sequencing data for the five variants (rs67052019, rs60911071, rs62558833, rs11108175, and rs9634162; Supplementary Table S3) that were highly significantly associated with colorectal cancer risk. Data for these five variants were extracted from whole-genome sequence datasets for 290 Shanghai colorectal cancer samples and compared with the genotypes derived from the imputation-based approach. The whole genome sequence was performed using the BGISEQ-500 sequencing platform with paired-end reads in length with 2 × 100 bp (mean read depth ~50M). The sequencing reads for each sample were mapped to the human reference genome (hg38) using the Burrows-Wheeler Aligner BWA program (version 0.75). The aligned reads were processed using the Genome Analysis ToolKit (GATKv3.7). Variant calling was performed individually for each sample with the GATK HaplotypeCaller tool and all samples together with GenotypeGVCFs to create a complete list of SNPs and indel VCFs. The Variant Quality Score Recalibration (VQSR) was then applied to filter variants of low quality.

Statistical analysis

We used the score test implemented in Rvtest (16) to associate genotype dosages with colorectal cancer risk after adjusting for age, sex, and the first five PCs of each individual study. SNPs with a low imputation quality ($R^2 < 0.3$) or a low MAF (<0.1% in the combined samples) were excluded from the downstream analysis. Summary statistics from each of 15 ACCC case–control studies were meta-analyzed using METAL (17) with the inverse variance-weighted fixed effect model. Associations with a $P < 5 \times 10^{-8}$ in the Asian studies alone or in combination with European studies were regarded as genome-wide significant. Each independent locus was defined as ±500 kb on either side of the most significant SNP that reached a genome-wide significant threshold ($P < 5 \times 10^{-8}$). The Cochran Q test (18) was used to evaluate the heterogeneity across studies and subgroups. We did not observe any apparent inflation in association statistics from the ACCC (Supplementary Fig. S2; refs. 6, 8, 11).

We performed approximate conditional analyses based on meta-analysis summary statistics to identify additional independent association signals at each locus using the GCTA-COJO method (19). The linkage disequilibrium (LD) matrix used in the analyses were based on 6,684 unrelated East Asian samples (interindividual genetic relationships < 0.025; ref. 11). Considering relatively small sample sizes of colorectal cancer studies in populations of East Asian ancestry (10, 11), variants that are conditionally independent of the index SNPs within a region and reached an empirical locus-wide significance of $P < 5 \times 10^{-5}$ were considered to be distinct association signals.

In silico functional characterization of novel loci

To identify potential functional variants and target genes at novel risk loci identified in this study, we annotated all variants that are in LD ($r^2 \geq 0.8$, EAS in the 1000 Genomes Project; 131 variants in total; Supplementary Table S4) with each of identified GWAS index variants within 500 kb using ANNOVAR (20). We used PolyPhen2 (21) and SIFT (22) to assess potential functional impact of coding variants. We further used regulatory information from the Roadmap Epigenomic project (epigenomic profiling) and the ENCODE project (regulatory protein binding and regulatory motifs) to characterize these 131 selected variants with the web-based HaploReg v4 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) as described previously (11). We tested for enrichment of colorectal cancer index variants with functional domains using the software of Genomic Regulatory Elements and Gwas Overlap algoRithm (GREGOR;ref. 23). This method tests for an increase in the number of colorectal cancer-associated index variants (94 independent GWAS index variants identified in populations of East Asian ancestry at $P < 5 \times 10^{-8}$ or replicated in populations of East Asian ancestry at $P < 5 \times 10^{-2}$ for those variants originally identified in populations of European ancestry; Supplementary Table S5), or their LD proxies (r^2 0.7, EAS in the 1000 Genomes Project), overlapping with the regulatory features (peaks from H3K4me1 and H3K27ac as enhancers, peaks from H3K4me3 and H3K9ac as promoters, and open chromatin as measured by DNase hypersensitivity) more often than expected by chance by comparing to permuted control sets in which the variants are matched on variant frequency, number of LD proxies, and distance to the nearest gene. A saddle-point approximation was used to estimate the P value based on the distribution of permuted statistics (23).

cis-expression quantitative trait loci analysis

We conducted *cis*-expression quantitative trait loci (*cis-e*QTLs) analyses for each of identified novel risk variants. The RNA was extracted from tumor-adjacent normal tissues obtained from 133 East Asian patients with colorectal cancer (7, 8, 11). We profiled gene expression using RNA sequencing with total mapped reads > 14M for each sample (11). We quantified gene expression levels using FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values. For *cis-eQTLs* analyses, we defined as a region ±1 mb within each risk variant. These patients were genotyped using the Illumina MEGA array as described before (11). We used linear regression models to associate gene expression levels with SNP genotypes with adjustment for sex and the top two PCs. We also evaluated associations of novel risk variants with gene expressions in 246 transverse colon tissues included in the Genotype-Tissue Expression (GTEx) database (24).

Results

In the meta-analysis of all Asian data (23,572 colorectal cancer cases and 48,700 controls; Supplementary Table S1), we identified an indel variant at 1p13.3 (rs67052019) with an OR of 1.13 (95% CI = 1.08–1.18) for colorectal cancer risk per deletion copy at the genome-wide significance level ($P < 5 \times 10^{-8}$; Table 1; Fig. 1; Supplementary Figs. S3 and S4). A variant (rs2938616) in complete LD with rs67052019 ($D'_{EAS\ or\ EUR} = 1.00$, $r^2_{EAS\ or\ EUR} = 1.00$) was associated with colorectal cancer risk in populations of European descent with OR of the deletion linked *G* allele = 1.03 (95% CI = 1.01− 1.04) and $P = 7.7\ 10^{-3}$ (Table 1), and this variant was also associated with colorectal cancer risk in populations of Asian descent [OR of the deletion linked *G* allele = 1.10 (95% CI = 1.05–1.15) and $P = 1.1 \times 10^{-5}$; Table 1].

Of the remaining 59 variants for replication in populations of European descent (58,131 colorectal cancer cases and 67,347 controls; Supplementary Table S2), 12 variants were replicated in the same direction as observed in East Asian populations at $P < 0.05$. Three variants (rs4308634, rs11108175, and rs9634162) were identified to be associated with colorectal cancer risk at the genome-wide significance level ($P < 5 \times 10^{-8}$) in the combined meta-analysis results of East-Asians and Europeans (Table 1; Fig. 1; Supplementary Figs. S3 and S4). However, the variant of rs4308634 [$OR_{\text{for the G allele}} = 1.04$ (95% CI = 1.03–1.06) and $P = 6.5 \times 10^{-9}$ in the combined meta-analyses] at the 7p12.3 was in high LD ($r^2_{EAS} = 0.68$) with a nearby variants (rs10951878) recently identified in populations of European ancestry (13). In addition, two variants (rs60911071 and rs62558833) were associated with colorectal cancer risk near the genome-wide significance level ($P = 5.8 \times 10^{-8}$ and 7.5×10^{-8}, respectively; Table 1; Fig. 1; Supplementary Fig. S4). Little heterogeneity of the effect size was observed between these two populations for these variants (P ≥ 0.01, Table 1). Stratification analyses of these newly identified risk variants (rs67052019, rs60911071, rs62558833, rs11108175, and rs9634162) by tumor site (colon or rectum) did not identify any significant heterogeneity ($P > 0.05$; Supplementary Table S6).

The average imputation quality for these five highly significant variants (rs67052019, rs60911071, rs62558833, rs11108175, and rs9634162) for colorectal cancer risk ranged from 0.78 to 0.99 in Asian cohorts and from 0.98 to 0.99 in European cohorts

(Supplementary Table S3). The overall genotype concordance rates for these five common variants (minor allele frequency ≥0.18; Table 1) between imputed data and whole-genome sequencing data were high (>0.90) based on 290 Shanghai colorectal cancer samples, indicating that the imputation quality was excellent for these SNPs (Supplementary Table S3).

Evaluation of local secondary signals

Independent secondary signals were reported in our previous studies for four colorectal cancer risk loci in East-Asian populations (*EIF3H* at 8q23.3, *NKX2–3* at 10q24.2, *VTI1A*/*TCF7L2* at 10q25.2 and *SMAD7* at 18q21.1; refs. 7, 8, 11, 14, 25). In this study, we observed 13 independent secondary signals at 11 additional loci in East Asians at the locus-wide significance level ($P < 5.0 \times 10^{-5}$; *SLCO2A1* at 3q22.2, *PITX1* at 5q31.1, *NOTCH4/HLA-DRB1/HLA-DRB5* at 6p21.32, *GATA3* at 10p14, *PRICKLE1* at 12q12, *LRP1/ARHGAP9* at 12q13.3, *MYL2/SH2B3* at 12q24.11, *WWOX/MAF* at 16q23.2, *NXN* at 17p13.3, *MYHAS/TMEM238L* at 17p12 and *BMP2* at 20p12.3; Table 2).

Functional characterization of risk loci and cis-eQTL analyses

We used functional genomics data to annotate each of the identified five novel variants (Table 1) as well as their correlated variants ($r^2_{EAS} \geq 0.80$).Aligning these risk variants with histone methylation/acetylation marks and DNase hypersensitivity sites (26) revealed that variants at three loci (1p13.3, 8p21.2, and 12q22) overlapped with the promoter/enhancer histone marks or DNase hypersensitivity sites in gastrointestinal tissues (Supplementary Table S4). This suggests that these variants may be involved in regulating gene expressions in gastrointestinal tissues. We further tested for tissue regulatory element enrichments of 94 colorectal cancer–associated variants identified or replicated in populations of East Asian ancestry (Supplementary Table S5 and Materials and Methods) using GREGOR (23). We found that colorectal cancer-associated variants were strongly associated with regulatory elements in gastrointestinal tissues (fetal stomach, fetal small intestine, fetal large intestine, small intestine, sigmoid colon, colonic mucosa, rectal mucosa, and stomach mucosa; >2.0 × enrichment; Supplementary Tables S7 and S8). Interestingly, monocytes were also enriched as indicated by DNase I hypersensitive sites ($P = 8.1\ 10^{-13}$, 2.3 × enrichment; Supplementary Tables S7 and S8). Similar results were observed on the basis of colorectal cancer-associated variants identified separately in populations of East Asian ancestry (fetal stomach, fetal small intestine, fetal large intestine, sigmoid colon, rectal mucosa, stomach mucosa, and monocytes; >2.0 × enrichment and $P < 7.2 \times 10^{-6}$) or populations of European ancestry (fetal stomach, fetal small intestine, fetal large intestine, sigmoid colon, colonic mucosa, and rectal mucosa; >2.0 × enrichment and $P < 6.7 \times 10^{-10}$).

We performed cis-eQTL analyses using transcriptome data from tumor-adjacent normal colon tissues from 133 patients with colorectal cancer of East Asian ancestry (Supplementary Table S9) and transverse colon tissues from 246 individuals predominantly of European ancestry in the GTEx (Supplementary Table S10). Significant correlations at $P < 0.05$ were found for 3 and 6 SNP–gene expression pairs in the East Asian and GTEx datasets, respectively. The colorectal cancer risk (deletion) allele of rs67052019 was associated with reduced expression of *UBL4B* and *GPR61*, but increased expression of

KCNC4-AS1 (consistent between the East Asian and GTEx datasets) and *TMEM167B*. The colorectal cancer risk T allele of rs62558833 was associated with increased expression of *SMU1*, *DCAF12*, and *NUDT2*. The CRC risk A allele of rs11108175 was associated with increased expression of *VEZT*.

Discussion

In this meta-analysis of 23,572 colorectal cancer cases and 48, 700 controls in East-Asians and follow-up replication analyses of 58,131 colorectal cancer cases and 67, 347 controls in individuals of European descent, we identified three novel risk loci and two highly suggestive loci for colorectal cancer risk. In addition, we identified 13 secondary signals at 11 known colorectal cancer risk loci in East Asians. Using functional genomics data, we showed that three of the newly identified risk variants, or their highly correlated variants, are located in regulatory regions of the genome. It indicates that these variants potentially regulate the expression of nearby genes in gastrointestinal tissues. In addition, monocytes were implicated in colorectal cancer carcinogenesis for the first time based on tissue regulatory element enrichment analyses. Our *cis*-eQTL analyses provide additional supports for a possible role of several risk variants identified inour study in regulating expression of cancer-related genes. Our study provides novel information toward the understanding of the genetic and biological basis of colorectal cancer.

At the 1p13.3 locus, the deletion allele of rs67052019 was associated with increased colorectal cancer risk. The variant rs67052019 was 58.8 Kb upstream of the *EPS8L3* gene. The function of the encoded protein by *EPS8L3* is unknown. Interestingly, the deletion allele of rs67052019 was associated with increased *KCNC4-AS1* (*KCNC4 antisense* RNA) expression in both Asian and GTEx datasets. The encode protein by *KCNC4* is the voltage gated Kv3.4 potassium channel protein involved in regulating mammalian cellcycle (27, 28). Theroles of Kv channels in cancer development and progression have been well established, and they are not only involved in cell proliferation and tumor growth, but also in cell migration and metastasis (29, 30).

At the 8p21.2 locus, rs60911071 is 35 kb downstream of *STC1*. Another variant (rs2928679, $r^2_{EAS\ or\ EUR} = 0.0$, between rs60911071 and rs2928679) in this region was reported to be associated with prostate cancer risk in populations of European ancestry (31). The encoded protein by *STC1* is a secreted glycoprotein, and is expressed ubiquitously, including in the gastrointestinal tract. STC-1 is reported to mediate the metastatic effect of platelet-derived growth factor signaling in colorectal cancer–associated fibroblasts (32). High expressions of *STC1* are correlated with poor postoperative survival in patients with colorectal cancer (33).

At the 9p13.3 locus, rs62558833 is an intronic variant of *UBAP2*. The encoded protein by *UBAP2* functions in the ubiquitination pathway. It inhibits the invasion of hepatocellular carcinoma cell by ubiquitinating and degrading Annexin A2 (34). The colorectal cancer risk *T* allele of rs62558833 was associated with an increased expression of *SMU1* (Supplementary Table S9), *DCAF12*, and *NUDT2* (Supplementary Table S10). The encoded protein by *SMU1* is suggested to be involved in genome stability maintenance by negatively regulating DNA synthesis (35). The encoded protein by *DCAF12* is required for

developmental apoptosis (36). The encoded protein by *NUDT2* is suggested to be a tumor-promoting factor, and high expressions of *NUDT2* are associated with poor prognosis and an increased risk of breast cancer recurrence (37, 38).

At the 12q22 locus, rs11108175 is a downstream variant of *NTN4*. Another nearby variant of rs17356907 ($r^2_{EAS\ or\ EUR}$ = 0.0,between rs11108175 and rs17356907) was reported to be associated with breast cancer risk in European populations (39). *NTN4* encodes a member of the netrin family of proteins that functions in various biological processes, including axonal guidance, tumorigenesis, and angiogenesis (40). NTN4 overexpression is observed to suppress primary and metastatic colorectal tumor progression through inhibiting tumor growth and angiogenesis (41–43). The colorectal cancer risk A allele of rs11108175 was nominally associated with an increased *VEZT* expression (Supplementary Table S10). The encode protein by *VEZT* is a ubiquitous transmembrane protein that is localized to adherens junctions in epithelial cells (44). How the aberrant expression of *VEZT* affects colorectal cancer risk warrants further investigations.

At the 12q24.21 locus, rs9634162 is 9.9 kb downstream of the *TBX3* gene. Two nearby variants, rs59336 ($P = 3.7 \times 10^{-7}$ on colorectal cancer risk, r^2_{EAS} = 0.58 and r^2_{EUR} = 0.66 between rs9634162 and rs59336) (45) and rs1427760 ($P = 2.5 \times 10^{-7}$ on colorectal cancer risk, r^2_{EAS} = 0.78 and r^2_{EUR} = 0.89 between rs9634162 and rs1427760; ref. 10) were reported to be associated with colorectal cancer risk in populations of European ancestry. We established this locus as one of the bona fide colorectal cancer risk loci at genome-wide significance. The protein encoded by *TBX3* belongs to the evolutionarily conserved T-box family of transcription factors that play critical roles in early embryonic development. Overexpression of Tbx3 is associated with multiple cancers potentially by modulating cell proliferation and survival, tumor formation and metastasis, and drug resistance (46). Emerging evidence suggests that Tbx3 is not only important for stem cell self-renewal, but also is extensively involved in cancer stemness (46, 47).

Independent secondary signals were observed at colorectal cancer risk loci in both East-Asian populations (7, 8, 11, 14, 25) and European populations (10, 13, 48–51). We observed 13 independent secondary signals at 11 additional colorectal cancer loci in East Asians. Conditional joint analysis could refine association signals and uncover additional GWAS loci for colorectal cancer. The loci of *MYL2/SH2B3* at 12q24.11 and *LRP1/ARHGAP9* at 12q13.3 that were reported in European populations for colorectal cancer risk (10, 52) only reached the genome-wide significance in the conditional joint analysis in East Asians. The missense variants of rs78894077 in *SH2B3* at 12q24.11 only exist in East Asian populations (Table 2) and is predicted to be highly “deleterious” in both PolyPhen2 (21) and SIFT (22). This strongly implicates that *SH2B3* is the underlying causal gene for colorectal cancer risk at this locus. The low-frequency intronic variant of rs368674461 in *MYHAS* at 17p12 also only exists in East Asian populations (Table 2); however, the function of *MYHAS* is currently unknown.

In summary, we identified three novel variants and two highly suggestive loci for colorectal cancer risk in this large GWAS of colorectal cancer. Combining information from functional annotations, *cis*-eQTL analyses and literature review, we propose the putative candidate

genes for these loci: *KCNC4-AS1* at 1p13.3, *STC1* at 8p21.2, *NUDT2* at 9p13.3, *NTN4* at 12q22, and *TBX3* at 12q24.21. Some of the putative target genes suggested by the results from our studies are located in established pathways for colorectal tumorigenesis, such as the maintaining of colon stem cells (*TBX3*). Multiple independent signals exist at many colorectal cancer loci; therefore, more familial relative risk of colorectal cancer could be explained at a locus when multiple independent signals were considered. However, extensive fine-mapping and functional follow-up studies are needed to identify the causal variants and target genes at each of the identified regions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Yingchang Lu[1], Sun-Seog Kweon[2,3], Qiuyin Cai[1], Chizu Tanikawa[4], Xiao-Ou Shu[1], Wei-Hua Jia[5], Yong-Bing Xiang[6], Jeroen R. Huyghe[7], Tabitha A. Harrison[7], Jeongseon Kim[8], Aesun Shin[9], Dong-Hyun Kim[10], Keitaro Matsuo[11,12], Sun Ha Jee[13], Xingyi Guo[1], Wanqing Wen[1], Jiajun Shi[1], Bingshan Li[14], Nan Wang[15], Min-Ho Shin[16], Hong-Lan Li[6], Zefang Ren[17], Jae Hwan Oh[18], Isao Oze[10], Yoon-Ok Ahn[19], Keum Ji Jung[20], Jing Gao[6], Yu-Tang Gao[6], Zhi-Zhong Pan[5], Yoichiro Kamatani[21,22], Andrew T. Chan[23,24,25,26], Andrea Gsur[27], Jochen Hampe[28], Loic Le Marchand[29], Li Li[30], Annika Lindblom[31,32], Victor Moreno[33,34,35], Polly A. Newcomb[7,36], Kenneth Offit[37,38], Paul D.P. Pharoah[39], Franzel J.B. van Duijnhoven[40], Bethany Van Guelpen[41], Pavel Vodicka[42,43,44], Stephanie J. Weinstein[45], Alicja Wolk[46], Anna H. Wu[47], Li Hsu[7,48], Yi-Xin Zeng[5], Jirong Long[1], Ulrike Peters[7,49], Koichi Matsuda[50], Wei Zheng[1]

Affiliations

[1]Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, Tennessee. [2]Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, South Korea. [3]Jeonnam Regional Cancer Center, Chonnam National University Hwasun Hospital, Hwasun, South Korea. [4]Laboratory of Genome Technology, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan. [5]State Key Laboratory of Oncology in South China, Cancer Center, Sun Yat-sen University, Guangzhou, China. [6]State Key Laboratory of Oncogene and Related Genes & Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China. [7]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington. [8]Department of Cancer Biomedical Science, Graduate School of Cancer Science and Policy, National Cancer Center, Gyeonggi-do, South Korea. [9]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul National University Cancer Research Institute, Seoul, Korea. [10]Department of Social and Preventive Medicine, Hallym University College of Medicine, Okcheon-dong, Korea. [11]Division of Molecular and

Clinical Epidemiology, Aichi Cancer Center Research Institute, Nagoya, Japan. [12]Department of Epidemiology, Nagoya University Graduate School of Medicine, Nagoya, Japan. [13]Department of Epidemiology and Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, Korea. [14]Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee. [15]General Surgery Department, Tangdu Hospital, Fourth Military Medical University, Xi'an, Shaanxi, China. [16]Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, South Korea. [17]School of Public Health, Sun Yat-sen University, Guangzhou, China. [18]Center for Colorectal Cancer, National Cancer Center Hospital, National Cancer Center, Gyeonggi-do, South Korea. [19]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, South Korea. [20]Institute for Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, Korea. [21]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Kanagawa, Japan. [22]Kyoto-McGill International Collaborative School in Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan. [23]Division of Gastroenterology, and Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. [24]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. [25]Broad Institute of Harvard and MIT, Cambridge, Massachusetts. [26]Department of Epidemiology, and Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts. [27]Institute of Cancer Research, Department of Medicine I, Medical University Vienna, Vienna, Austria. [28]Department of Medicine I, University Hospital Dresden, Technische Universität Dresden (TU Dresden), Dresden, Germany. [29]University of Hawaii Cancer Center, Honolulu, Hawaii. [30]Case Comprehensive Cancer Center, Case Western Reserve University, Cleveland, Ohio. [31]Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. [32]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [33]Cancer Prevention and Control Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain. [34]CIBER Epidemiología y Salud Publica (CIBERESP), Madrid, Spain. [35]Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain. [36]School of Public Health, University of Washington, Seattle, Washington. [37]Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York. [38]Department of Medicine, Weill Cornell Medical College, New York, New York. [39]Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom. [40]Division of Human Nutrition, Wageningen University and Research, Wageningen, the Netherlands. [41]Department of Radiation Sciences, Oncology Unit, Umeå University, Umeå, Sweden. [42]Department of Molecular Biology of Cancer, Institute of Experimental Medicine of the Czech Academy of Sciences, Prague, Czech Republic. [43]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, Czech Republic. [44]Faculty of Medicine and Biomedical Center in Pilsen, Charles University, Pilsen, Czech

Republic. [45]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland. [46]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [47]University of Southern California, Preventative Medicine, Los Angeles, California. [48]Department of Biostatistics, University of Washington, Seattle, Washington. [49]Department of Epidemiology, University of Washington, Seattle, Washington. [50]Laboratory of Clinical Genome Sequencing, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, University of Tokyo, Tokyo, Japan.

Acknowledgments

The authors thank all study participants and research staff of all parent studies for their contributions and commitment to this project. The authors thank Vanderbilt staff members Ms. Jing He for data processing and analyses and Mr. Marshal Younger for editing and preparing the manuscript. The work at Vanderbilt University Medical Center was supported by U.S. NIH grants R01CA188214, R37CA070867, UM1CA182910, R01CA124558, R01CA158473, and R01CA148667, as well as Anne Potter Wilson Chair funds from the Vanderbilt University School of Medicine. Sample preparation and genotyping assays at Vanderbilt University were conducted at the Survey and Biospecimen Shared Resources and Vanderbilt Microarray Shared Resource, which are supported in part by the Vanderbilt-Ingram Cancer Center (P30CA068485). Imputation and statistical analyses were performed on servers maintained by the Advanced Computing Center for Research and Education at Vanderbilt University (Nashville, TN). Studies (listed with grant support) participating in the Asia Colorectal Cancer Consortium include the Shanghai Women's Health Study (US NIH, R37CA070867, UM1CA182910), the Shanghai Men's Health Study (US NIH, R01CA082729, UM1CA173640), the Shanghai Breast and Endometrial Cancer Studies (US NIH, R01CA064277 and R01CA092585; contributing only controls), the Shanghai Colorectal Cancer Study 3 (US NIH, R37CA070867, R01CA188214 and Anne Potter Wilson Chair funds), the Guangzhou Colorectal Cancer Study (National Key Scientific and Technological Project, 2011ZX09307–001-04; the National Basic Research Program, 2011CB504303, contributing only controls, the Natural Science Foundation of China, 81072383, contributing only controls), the Hwasun Cancer Epidemiology Study–Colon and Rectum Cancer (HCES-CRC; grants from Chonnam National University Hwasun Hospital Biomedical Research Institute, HCRI18007), the Japan BioBank Colorectal Cancer Study (grant from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government), the Aichi Colorectal Cancer Study(Grant-in-Aid for Cancer Research, grant for the Third Term Comprehensive Control Research for Cancer and Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology, 17015018 and 221S0001), the Korea-NCC (National Cancer Center) Colorectal Cancer Study (Basic Science Research Program through the National Research FoundationofKorea, 2010–0010276and 2013R1A1A2A10008260; National Cancer Center Korea, 0910220), and the KCPS-II Colorectal Cancer Study (National R&D Program for Cancer Control, 1631020; Seoul R&D Program, 10526). Funding information for the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO) and its participating studies is provided in this section. GECCO: NCI, NIH, U.S. Department of Health and Human Services (U01 CA164930, U01 CA137088, R01 CA059045, U01 CA164930, R21 CA191312, R01201407). Genotyping/ Sequencing services were provided by the Center for Inherited Disease Research (CIDR; X01-HG008596 and X-01-HG007585). CIDR is fully funded through a federal contract from the NIH to The Johns Hopkins University, contract number HHSN268201200008I. This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA015704. ASTERISK: a Hospital Clinical Research Program (PHRC-BRD09/C) from the University Hospital Center of Nantes (CHU de Nantes) and supported by the Regional Council of Pays de la Loire, the Groupement des Entreprises Francaises dans la Lutte contre le Cancer (GEFLUC), the Association Anne de Bretagne Génétique and the Ligue Régionale Contre le Cancer (LRCC). The ATBC Study is supported by the Intramural Research Program of the U.S. NCI, NIH, and by U.S. Public Health Service contract HHSN261201500005C from the NCI,DepartmentofHealth andHuman Services. CLUE II: Thisresearchwas funded by the American Institute for Cancer Research and the Maryland Cigarette Restitution Fund at Johns Hopkins, and the NCI (P30 CA006973, to W.G. Nelson). COLO2&3: NIH (R01 CA60987). ColoCare: This work was supported by the NIH [grant numbers R01 CA189184 (Li/Ulrich), U01 CA206110 (Ulrich/Li/ Siegel/Figueireido/Colditz, 2P30CA015704– 40 (Gilliland), R01 CA207371 (Ulrich/Li)], the Matthias Lackas-Foundation, the German Consortium for Translational Cancer Research, and the EU TRANSCAN initiative. The Colon Cancer Family Registry (CFR) Illumina GWAS was supported by funding from the NCI, NIH (grant numbers U01 CA122839, R01 CA143247, to G. Casey). The Colon CFR participant recruitment and collection of data and biospecimens used in this study were supported by the NCI, NIH (grant number U01 CA167551). The content of this manuscript does not necessarily reflect the views or policies of the NCI or any of the collaborating centers in the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government, any cancer registry, or the CCFR. COLON: The COLON study is sponsored by Wereld Kanker Onderzoek Fonds, including funds from grant 2014/ 1179 as part of

the World Cancer Research Fund International Regular Grant Programme, by Alpe d'Huzes and the Dutch Cancer Society (UM 2012–5653, UW 2013–5927, UW2015–7946), and by TRANSCAN (JTC2012-MetaboCCC, JTC2013-FOCUS). The Nqplus study is sponsored by a ZonMW investment grant (98–10030); by PREVIEW, the project PREVention of diabetes through lifestyle intervention and population studies in Europe and around the World (PREVIEW) project which received funding from the European Union Seventh Framework Programme (FP7/ 2007–2013) under grant no. 312057; by funds from TI Food and Nutrition (cardiovascular health theme), a public–private partnership on precompetitive research in food and nutrition; and by FOODBALL, the Food Biomarker Alliance, a project from JPI Healthy Diet for a Healthy Life. Colorectal Cancer Transdisciplinary (CORECT) Study: The CORECT Study was supported by the NCI/NIH, U.S. Department of Health and Human Services (grant numbers U19 CA148107, R01 CA81488, P30 CA014089, R01 CA197350; P01 CA196569; R01 CA201407) and National Institutes of Environmental Health Sciences, National Institutes of Health (grant number T32 ES013678). CORSA: “Österreichische Nationalbank Jubiläumsfondsprojekt” (12511) and Austrian Research Funding Agency (FFG) grant 829675. CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study-II (CPS-II) cohort. This study was conducted with Institutional Review Board approval. CRCGEN: Colorectal Cancer Genetics & Genomics, Spanish study was supported by Instituto de Salud Carlos III, co-funded by FEDER funds –a way to build Europe– (grants PI14–613 and PI09–1286), Agency for Management of University and Research Grants (AGAUR) of the Catalan Government (grant 2017SGR723), and Junta de Castilla y León (grant LE22A10–2). Sample collection of this work was supported by the Xarxa de Bancs de Tumors de Catalunya sponsored by Pla Director d'Oncología de Catalunya (XBTC), Plataforma Biobancos PT13/0010/0013 and ICOBIOBANC, sponsored by the Catalan Institute of Oncology. Czech Republic CCS: This work was supported by the Grant Agency of the Czech Republic (grants CZ GA CR: GAP304/10/1286 and 1585) and by the Grant Agency of the Ministry of Health of the Czech Republic (grants AZV 15–27580A and AZV 17–30920A). DACHS: This work was supported by the German Research Council (BR 1704/6–1, BR 1704/6–3, BR 1704/6–4, CH 117/1–1, HO 5117/2–1, HE 5998/2–1, KL 2354/3–1, RO 2270/8–1 and BR 1704/17–1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B). DALS: NIH (R01 CA48998, to M.L. Slattery). EDRN: This work is funded and supported by the NCI, EDRN Grant (U01 CA 84968–06). EPIC: The coordination of EPIC is financially supported by the European Commission (DGSANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Généralede l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM;France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRCItaly and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/ 01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020; Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/ A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/ M012190/1 to EPICOxford; United Kingdom). EPICOLON: This work was supported by grants from Fondo de Investigación Sanitaria/FEDER (PI08/0024, PI08/ 1276, PS09/02368, P111/00219, PI11/00681, PI14/00173, PI14/00230, PI17/00509, 17/00878, Acción Transversal de Cancer), Xunta de Galicia (PGIDIT07P-XIB9101209PR), Ministerio de Economia y Competitividad (SAF07–64873, SAF 2010–19273, SAF2014–54453R), Fundación Científica de la Asociación Española contra el Cáncer (GCB13131592CAST), Beca Grupode Trabajo “Oncología” AEG (Asociación Española de Gastroenterología), Fundación Privada Olga Torres, FP7 CHIBCHA Consortium, Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya, 2014SGR135, 2014SGR255, 2017SGR21, 2017SGR653), Catalan Tumour Bank Network (Pla Director d'Oncologia, Generalitat de Catalunya), PERIS (SLT002/16/00398, Generalitat de Catalunya), CERCA Programme (Generalitat de Catalunya) and COST Action BM1206. CIBERehd is funded by the Instituto de Salud Carlos III. ESTHER/VERDI: This work was also supported by grants from the Baden-Württemberg Ministry of Science, Research and Arts and the German Cancer Aid. Harvard cohorts (HPFS, NHS, PHS): The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required. HPFS is supported by the NIH (P01 CA055075, UM1 CA167552, U01 CA167552, R01 CA137178, R01 CA151993, R35 CA197735, K07 CA190673, and P50 CA127003), NHS by the NIH (R01 CA137178, P01 CA087969, UM1 CA186107, R01 CA151993, R35 CA197735, K07 CA190673, and P50 CA127003), and PHS by the NIH (R01 CA042182). Hawaii Adenoma Study: NCI grants R01 CA72520. HCES-CRC: the Hwasun Cancer Epidemiology Study–Colon and Rectum Cancer (HCES-CRC; grants from Chonnam National University Hwasun Hospital, HCRI15011–1). Kentucky: This work was supported by the following grant support: Clinical Investigator Award from Damon Runyon Cancer Research Foundation (CI-8); NCI R01CA136726. LCCS: The Leeds Colorectal Cancer Study was funded by the Food Standards Agency and Cancer Research UK Programme Award (C588/A19167). MCCScohortrecruitmentwasfundedbyVicHealthandCancerCouncilVictoria. The MCCS was further supported by Australian NHMRC grants 509348, 209057, 251553, and 504711 and by infrastructure provided by Cancer Council

Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database. MEC: NIH (R37 CA54281, P01 CA033619, and R01 CA063464). MECC: This work was supported by the NIH, U.S. Department of Health and Human Services (R01 CA81488, to S.B. Gruber and G. Rennert). MSKCC: The work at Sloan Kettering in New York was supported by the Robert and Kate Niehaus Center for Inherited Cancer Genomics and the Romeo Milio Foundation. Moffitt: This work was supported by funding from the NIH (grant numbers R01 CA189184, P30 CA076292), Florida Department of Health Bankhead-Coley Grant 09BN-13, and the University of South Florida Oehler Foundation. Moffitt contributions were supported in part by the Total Cancer Care Initiative, Collaborative Data Services Core, and Tissue Core at the H. Lee Moffitt Cancer Center & Research Institute, a NCI-designated Comprehensive Cancer Center (grant number P30 CA076292). NCCCS I & II: We acknowledge funding support for this project from the NIH, R01 CA66635 and P30 DK034987. NFCCR: This work was supported by an Interdisciplinary Health Research Team award from the Canadian Institutes of Health Research (CRT 43821); the NIH, U.S. Department of Health and Human Serivces (U01 CA74783); and National Cancer Institute of Canada grants (18223 and 18226). The authors wish to acknowledge the contribution of Alexandre Belisle and the genotyping team of the McGill University and Génome Québec Innovation Centre, Montréal, Canada, for genotyping the Sequenom panel in the NFCCR samples. Funding was provided to Michael O. Woods by the Canadian Cancer Society Research Institute. NSHDS: Swedish Research Council; the Swedish Cancer Society; Region Västerbotten; the Lion's Cancer Research Foundation, the Faculty of Medicine and Insamlingsstiftelsen, all at Umeå University; and the Margareta Dannborg Memorial Fund. OFCCR: NIH, through funding allocated to the Ontario Registry for Studies of Familial Colorectal Cancer (U01 CA074783); see CCFR section above. Additional funding toward genetic analyses of OFCCR includes the Ontario Research Fund, the Canadian Institutes of Health Research, and the Ontario Institute for Cancer Research, through generous support from the Ontario Ministry of Research and Innovation. OSUMC: OCCPI funding was provided by Pelotonia, and HNPCC funding was provided by the NCI (CA16058 and CA67941). PLCO: Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS. Funding was provided by NIH, Genes, Environment and Health Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438. PMH: NIH (R01 CA076366, to P.A. Newcomb). SEARCH: The University of Cambridge has received salary support in respect of PDPP from the NHS in the East of England through the Clinical Academic Reserve. Cancer Research UK (C490/A16561); the UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge. SELECT: The Selenium and Vitamin E Cancer Prevention Trial (SELECT) was supported by the NCI of the NIH under award numbers UM1CA182883 and U10CA37429. SMS: This work was supported by the National Cancer Institute (grant P01 CA074184, to J.D. Potter and P.A. Newcomb; grants R01 CA097325, R03 CA153323, and K05 CA152715, to P.A. Newcomb), and the National Center for Advancing Translational Sciences at the NIH (grant KL2 TR000421, to A.N. Burnett-Hartman) The Swedish Low-risk Colorectal Cancer Study: The study was supported by grants from the Swedish Research Council; K2015–55X-22674–01-4, K2008–55X-20157–03-3, K2006–72X-20157–01-2, and the Stockholm County Council (ALF project). Swedish Mammography Cohort and Cohort of Swedish Men: This work is supported by the Swedish Research Council/Infrastructure grant, the Swedish Cancer Foundation, and the Karolinska Institute's Distinguished Professor Award to A. Wolk. UK Biobank: This research has been conducted using the UK Biobank Resource under Application Number 8614VITAL: NIH (K05 CA154337). WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, NIH, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. ASTERISK: We are very grateful to Dr. Bruno Buecher without whom this project would not have existed. CLUE II: We thank Judith Hoffman-Bolton, Senior Research Program Coordinator, for her contributions to the conduct of CLUE. CORSA: We are grateful to Doris Mejri and Monika Hunjadi for laboratory assistance. CPS-II: Theauthors would like to acknowledge the contribution to this study from central cancer registries supported through the CDC and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance, Epidemiology, and End Results Program. DACHS: Ute Handte-Daub, Utz Benscheid, Muhabbet Celik, and Ursula Eilber for excellent technical assistance. EDRN: We acknowledge the following contributors to the development of the resource: University of Pittsburgh School of Medicine, Department of Gastroenterology, Hepatology and Nutrition: Lynda Dzubinski; University of Pittsburgh School of Medicine, Department of Pathology: Michelle Bisceglia; and University of Pittsburgh School of Medicine, Department of Biomedical Informatics. EPICOLON: We acknowledge the Spanish National DNA Bank, Biobank of Hospital Clínic–IDIBAPS and Biobanco Vasco for the availability of the samples. The work was carried out (in part) at the Esther Koplowitz Centre, Barcelona. Harvard cohorts (HPFS, NHS, PHS): We would like to thank the participants and staff of the HPFS, NHS, and PHS for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data. LCCS: We acknowledge the contributions of Jennifer Barrett, Robin Waxman, Gillian Smith,and Emma Northwoodin conducting this study. NSHDS: We thank the Biobank Research Unit at Umea University, Biobanken Norr at Region Västerbotten, and the cohort participants. WHI: The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full list of WHI investigators can be found at http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf. Collectively, the authors thank all those who helped make this research possible, including patients, healthy control persons, physicians, staff, technicians,

investigators, students, participating clinics and hospitals, state registries, and study teams for their participation in this study.

References

1. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the Global Burden of Disease Study. JAMA Oncol 2017;3:524–48. [PubMed: 27918777]
2. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial risk and heritability of cancer among twins in Nordic countries. JAMA 2016;315:68–76. [PubMed: 26746459]
3. Peters U, Bien S, Zubair N. Genetic architecture of colorectal cancer. Gut 2015; 64:1623–36. [PubMed: 26187503]
4. Adam R, Spier I, Zhao B, Kloth M, Marquez J, Hinrichsen I, et al. Exome sequencing identifies Biallelic MSH3 germline mutations as a recessive subtype of colorectal adenomatous polyposis. Am J Hum Genet 2016;99:337–51. [PubMed: 27476653]
5. Weren RD, Ligtenberg MJ, Kets CM, de Voer RM, Verwiel ET, Spruijt L, et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. Nat Genet 2015;47:668–71. [PubMed: 25938944]
6. Jia WH, Zhang B, Matsuo K, Shin A, Xiang YB, Jee SH, et al. Genome-wide association analyses in East Asians identify new susceptibility loci for colorectal cancer. Nat Genet 2013;45:191–6. [PubMed: 23263487]
7. Zeng C, Matsuda K, Jia WH, Chang J, Kweon SS, Xiang YB, et al. Identification of susceptibility loci and genes for colorectal cancer risk. Gastroenterology 2016; 150:1633–45. [PubMed: 26965516]
8. Zhang B, Jia WH, Matsuda K, Kweon SS, Matsuo K, Xiang YB, et al. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. Nat Genet 2014;46:533–42. [PubMed: 24836286]
9. Schmit SL, Edlund CK, Schumacher FR, Gong J, Harrison TA, Huyghe JR, et al. Novel common genetic susceptibility loci for colorectal cancer. J Natl Cancer Inst 2019;111:146–57. [PubMed: 29917119]
10. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet 2019; 51:76–87. [PubMed: 30510241]
11. Lu Y, Kweon SS, Tanikawa C, Jia WH, Xiang YB, Cai Q, et al. Large-scale genome-wide association study of East Asians identifies loci associated with risk for colorectal cancer. Gastroenterology 2019;156:1455–66. [PubMed: 30529582]
12. Dunlop MG, Dobbins SE, Farrington SM, Jones AM, Palles C, Whiffin N, et al. Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. Nat Genet 2012;44:770–6. [PubMed: 22634755]
13. Law PJ, Timofeeva M, Fernandez-Rozadilla C, Broderick P, Studd J, Fernandez-Tajes J, et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. Nat Commun 2019;10:2154. [PubMed: 31089142]
14. Zhang B, Jia WH, Matsuo K, Shin A, Xiang YB, Matsuda K, et al. Genome-wide association study identifies a new SMAD7 risk variant associated with colorectal cancer risk in East Asians. Int J Cancer 2014;135:948–55. [PubMed: 24448986]
15. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–9. [PubMed: 16862161]
16. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics 2016;32:1423–6. [PubMed: 27153000]

17. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–1. [PubMed: 20616382]
18. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. Ann Intern Med 1997;127:820–6. [PubMed: 9382404]
19. Yang J, Ferreira T, Morris AP, Medland SE; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012;44:369–75. [PubMed: 22426310]
20. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38: e164. [PubMed: 20601685]
21. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;Chapter 7: Unit7.20.
22. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–81. [PubMed: 19561590]
23. Schmidt EM, Zhang J, Zhou W, Chen J, Mohlke KL, Chen YE, et al. GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. Bioinformatics 2015;31:2601–6. [PubMed: 25886982]
24. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648–60. [PubMed: 25954001]
25. Wang H, Burnett T, Kono S, Haiman CA, Iwasaki M, Wilkens LR, et al. Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A. Nat Commun 2014;5:4613. [PubMed: 25105248]
26. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res 2016;44:D877–81. [PubMed: 26657631]
27. Kunzelmann K. Ion channels and cancer. J Membr Biol 2005;205:159–73. [PubMed: 16362504]
28. Leblanc N. Kv3.4, a key signalling molecule controlling the cell cycle and proliferation of human arterial smooth muscle cells. Cardiovasc Res 2010;86: 351–2. [PubMed: 20378681]
29. Huang X, Jan LY. Targeting potassium channels in cancer. J Cell Biol 2014;206: 151–62. [PubMed: 25049269]
30. Pardo LA, Stuhmer W. The roles of K(+) channels in cancer. Nat Rev Cancer 2014;14:39–48. [PubMed: 24336491]
31. Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116–21. [PubMed: 19767753]
32. Pena C, Cespedes MV, Lindh MB, Kiflemariam S, Mezheyeuski A, Edqvist PH, et al. STC1 expression by cancer-associated fibroblasts drives metastasis of colorectal cancer. Cancer Res 2013;73:1287–97. [PubMed: 23243022]
33. Tamura S, Oshima T, Yoshihara K, Kanazawa A, Yamada T, Inagaki D, et al. Clinical significance of STC1 gene expression in patients with colorectal cancer. Anticancer Res 2011;31:325–9. [PubMed: 21273618]
34. Bai DS, Wu C, Yang LX, Zhang C, Zhang PF, He YZ, et al. UBAP2 negatively regulates the invasion of hepatocellular carcinoma cell by ubiquitinating and degradating Annexin A2. Oncotarget 2016;7:32946–55. [PubMed: 27121050]
35. Ren L, Liu Y, Guo L, Wang H, Ma L, Zeng M, et al. Loss of Smu1 function derepresses DNA replication and over-activates ATR-dependent replication checkpoint. Biochem Biophys Res Commun 2013;436:192–8. [PubMed: 23727573]
36. Hwangbo DS, Biteau B, Rath S, Kim J, Jasper H. Control of apoptosis by Drosophila DCAF12. Dev Biol 2016;413:50–9. [PubMed: 26972874]
37. Kwon O, Kwak D, Ha SH, Jeon H, Park M, Chang Y, et al. Nudix-type motif 2 contributes to cancer proliferation through the regulation of Rag GTPase-mediated mammalian target of rapamycin complex 1 localization. Cell Signal 2017;32:24–35. [PubMed: 28089905]

38. Oka K, Suzuki T, Onodera Y, Miki Y, Takagi K, Nagasaki S, et al. Nudix-type motif 2 in human breast carcinoma: a potent prognostic factor associated with cell proliferation. Int J Cancer 2011;128:1770–82. [PubMed: 20533549]
39. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 2013;45:353–61. [PubMed: 23535729]
40. Wilson BD, Ii M, Park KW, Suli A, Sorensen LK, Larrieu-Lahargue F, et al. Netrins promote developmental and therapeutic angiogenesis. Science 2006;313: 640–4. [PubMed: 16809490]
41. Eveno C, Broqueres-You D, Feron JG, Rampanou A, Tijeras-Raballand A, Ropert S, et al. Netrin-4 delays colorectal cancer carcinomatosis by inhibiting tumor angiogenesis. Am J Pathol 2011;178:1861–9. [PubMed: 21406174]
42. Eveno C, Contreres JO, Hainaud P, Nemeth J, Dupuy E, Pocard M. Netrin-4 overexpression suppresses primary and metastatic colorectal tumor progression. Oncol Rep 2013;29:73–8. [PubMed: 23124703]
43. Lejmi E, Leconte L, Pedron-Mazoyer S, Ropert S, Raoul W, Lavalette S, et al. Netrin-4 inhibits angiogenesis via binding to neogenin and recruitment of Unc5B. Proc Natl Acad Sci U S A 2008;105:12491–6. [PubMed: 18719102]
44. Kussel-Andermann P, El-Amraoui A, Safieddine S, Nouaille S, Perfettini I, Lecuit M, et al. Vezatin, a novel transmembrane protein, bridges myosin VIIA to the cadherin-catenins complex. EMBO J 2000;19:6020–9. [PubMed: 11080149]
45. Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, et al. Identification of genetic susceptibility loci for colorectal tumors in a genomewide meta-analysis. Gastroenterology 2013;144:799–807. [PubMed: 23266556]
46. Dong L,Lyu X,Faleti OD,He ML.Thespecial stemness functions ofTbx3 instem cells and cancer development. Semin Cancer Biol 2019;57:105–10. [PubMed: 30268432]
47. Russell R, Ilg M, Lin Q, Wu G, Lechel A, Bergmann W, et al. A dynamic role of TBX3 in the pluripotency circuitry. Stem Cell Reports 2015;5:1155–70. [PubMed: 26651606]
48. Whiffin N, Hosking FJ, Farrington SM, Palles C, Dobbins SE, Zgaga L, et al. Identification of susceptibility loci for colorectal cancer in a genome-wide metaanalysis. Hum Mol Genet 2014;23:4729–37. [PubMed: 24737748]
49. Tomlinson IP, Webb E, Carvajal-Carmona L, Broderick P, Howarth K, Pittman AM, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat Genet 2008;40:623–30. [PubMed: 18372905]
50. Study C, Houlston RS, Webb E, Broderick P, Pittman AM, Di Bernardo MC, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nat Genet 2008;40:1426–35. [PubMed: 19011631]
51. Tomlinson IP, Carvajal-Carmona LG, Dobbins SE, Tenesa A, Jones AM, Howarth K, et al. Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer. PLoS Genet 2011;7:e1002105. [PubMed: 21655089]
52. Schumacher FR, Schmit SL, Jiao S, Edlund CK, Wang H, Zhang B, et al. Genomewide association study of colorectal cancer identifies six new susceptibility loci. Nat Commun 2015;6:7138. [PubMed: 26151821]

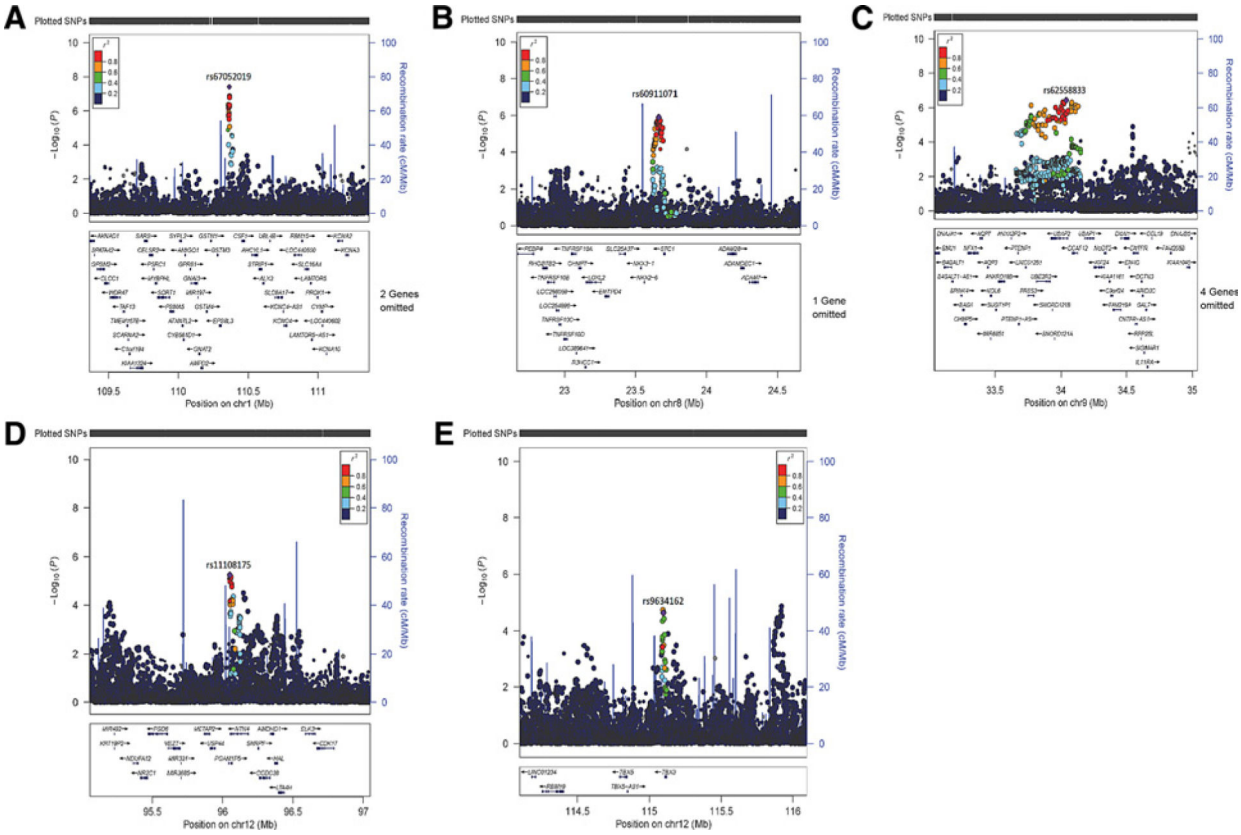

Figure 1.

Regional association plots of five colorectal cancer risk loci in Asian descendants (**A**, rs67052019; **B**, rs60911071; **C**, rs62558833; **D**, rs11108175; and **E**, rs9634162). Each dot represents the *P* value (on a $\log_{10}$ scale) of an SNP with colorectal cancer risk based on the meta-analysis results in East Asians only and is presented according to its genomic position (NCBI Build 37). The most significantly associated SNP in the combined meta-analyses is represented by purple. The color of all other SNPs indicates the level of LD with the lead SNP (estimated by ASN r^2 from the 1000 Genome Project data). Recombination rates were also estimated from 1000 Genomes Project data, and gene annotations within the 2 Mb regions that are centered on the newly identified risk variants were obtained from the UCSC Genome Browser.

Table 1.

Novel loci associated with colorectal cancer risk identified in this study.

					East Asian descendants (up to 23,572 cases and 48,700 controls)				European descendants (up to 58,131 cases and 67,347 controls)				Combined meta-analysis		
Locus	**SNP**	**Position (hg19)**	**Nearby genes**	**Alleles**	**RAF**	**OR (95% CI)**	**P**	**P_{het}**	**RAF**	**OR (95% CI)**	**P**	**P_{het}**	**OR (95% CI)**	**P**	**P_{het}**
Loci reaching genome-wide significance with $P < 5 \times 10^{-8}$															
1p13.3	rs67052019	110,365,461	EPS8L3	Del [a]/Ref	0.18	1.13 (1.08–1.18)	3.9×10^{-8}	0.28	–	NA [b]	–	–	–	–	–
12q22	rs11108175	96,050,887	NTN4	A/G	0.80	1.08 (1.05–1.12)	5.7×10^{-6}	0.10	0.59	1.05 (1.03–1.07)	2.6×10^{-7}	0.88	1.05 (1.04–1.07)	2.7×10^{-11}	0.10
12q24.21	rs9634162	115,098,094	TBX3	A/G	0.48	1.07 (1.03–1.10)	2.4×10^{-5}	0.68	0.51	1.04 (1.03–1.06)	7.5×10^{-7}	0.02	1.05 (1.03–1.07)	2.3×10^{-10}	0.24
Loci near genome-wide significance with $P < 1 \times 10^{-7}$ but $> 5 \times 10^{-8}$															
8p21.2	rs60911071	23,664,632	STC1	G/C	0.64	1.07 (1.04–1.10)	1.2×10^{-6}	0.8	0.95	1.05 (1.01–1.10)	3.2×10^{-2}	0.14	1.07 (1.04–1.09)	5.8×10^{-8}	0.48
9p13.3	rs62558833	34,039,002	UBAP2	T/C	0.55	1.08 (1.05–1.11)	3.6×10^{-7}	0.1	0.29	1.03 (1.01–1.05)	1.7×10^{-3}	0.62	1.05 (1.03–1.06)	7.5×10^{-8}	0.01

Abbreviations: alleles, risk allele/reference allele; CI, confidence interval; P_{het}: P value derived from the heterogeneity test; RAF, risk allele frequency.

[a] Del: deletion for ACAGAGAGATGTAGGGGC.

[b] Variant of rs2938616 ($D'_{EAS\ or\ EUR} = 1.00$, $r^2_{EAS\ or\ EUR} = 1.00$ with rs67052019) was associated with colorectal cancer risk in populations of European ancestry [OR for the deletion linked G allele of rs2938616 = 1.03 (95% CI = 1.01–1.04) and $P = 7.7 \times 10^{-3}$] and in populations of East Asian ancestry [OR of the deletion linked G allele of rs2938616 = 1.10 (95% CI = 1.05–1.15) and $P = 1.1 \times 10^{-5}$].

Table 2.

Secondary independent association signals ($P < 5 \times 10^{-5}$) identified within 500 kb of previously reported risk loci for colorectal cancer: results from the ACCC.

						Single-SNP meta-analysis			Joint analysis			Multiple-independent-variant loci New/known[b]
Locus	SNP	CHR	Pos	Nearest gene	EA	Freq	OR (95% CI)	P	OR (95% CI)	P	r^a	
3q22.2	rs4854776	3	133,687,864	SLCO2A1	A	0.78	0.94 (0.91–0.97)	2.43×10^{-5}	0.93 (0.91–0.96)	1.28×10^{-5}	−0.04	New
	rs61510274	3	133,749,515		C	0.41	0.90 (0.88–0.93)	2.68×10^{-14}	0.90 (0.88–0.93)	2.77×10^{-13}	−0.11	
	rs76941686	3	133,815,683		C	0.65	1.07 (1.05–1.10)	1.60×10^{-7}	1.06 (1.03–1.09)	1.78×10^{-5}		
5q31.1	rs7722513	5	134,464,066	PITX1	C	0.25	1.15 (1.11–1.18)	9.14×10^{-20}	1.16 (1.12–1.19)	7.30×10^{-22}	−0.08	New
	rs35917784	5	134,497,599		A	0.10	1.10 (1.05–1.15)	6.56×10^{-5}	1.12 (1.07–1.17)	4.09×10^{-6}		
6p21.32	rs3830041	6	32,191,339	NOTCH4/HLA-DRB1/HLA-DRB5	T	0.14	1.15 (1.10–1.21)	3.61×10^{-9}	1.16 (1.11–1.22)	1.92×10^{-9}	−0.09	New
	rs4713534	6	32,445,926		T	0.07	1.14 (1.06–1.23)	2.54×10^{-4}	1.17 (1.09–1.26)	1.04×10^{-5}	−0.08	
	rs569582972	6	32,558,756		T	0.05	1.21 (1.10–1.33)	7.74×10^{-5}	1.23 (1.12–1.36)	1.88×10^{-5}		
8q23.3	rs6469654	8	117,632,965	EIF3H	C	0.50	1.12 (1.09–1.15)	1.72×10^{-17}	1.11 (1.08–1.14)	3.49×10^{-16}	0.09	Known
	rs12541711	8	117,715,457		A	0.03	1.26 (1.15–1.37)	7.67×10^{-7}	1.22 (1.11–1.34)	1.65×10^{-5}		
10p14	rs533919	10	8,266,151	GATA3	A	0.08	0.90 (0.86–0.95)	1.91×10^{-5}	0.91 (0.86–0.95)	2.54×10^{-5}	−0.00	New
	rs7894531	10	8,734,761		A	0.36	0.86 (0.84–0.89)	2.52×10^{-28}	0.86 (0.84–0.89)	2.24×10^{-28}		
10q24.2	rs6584283	10	101,290,301	NKX2–3, SLC25A28	T	0.44	0.92 (0.90–0.94)	6.02×10^{-11}	0.92 (0.90–0.95)	3.07×10^{-10}	−0.05	Known
	rs61871279	10	101,343,705		T	0.17	1.08 (1.05–1.12)	5.37×10^{-6}	1.08 (1.04–1.11)	1.96×10^{-5}		
10q25.2	rs4554811	10	114,278,734	VTI1A/TCF7L2	A	0.73	0.90 (0.87–0.92)	8.08×10^{-14}	0.90 (0.87–0.93)	4.08×10^{-13}	−0.03	Known
	rs11196172	10	114,726,843		A	0.69	1.12 (1.09–1.15)	1.74×10^{-15}	1.12 (1.09–1.15)	6.53×10^{-15}		
12q12	rs117912059	12	43,006,211	PRICKLE1	T	0.98	1.23 (1.12–1.36)	3.76×10^{-5}	1.23 (1.11–1.36)	4.26×10^{-5}	−0.02	New
	rs2730985	12	43,130,624		A	0.37	0.92 (0.90–0.95)	1.24×10^{-9}	0.92 (0.90–0.95)	1.34×10^{-9}		
12q13.3	rs7398375	12	57,540,848	LRP1/ARHGAP9	C	0.43	1.08 (1.05–1.12)	1.27×10^{-7}	1.09 (1.06–1.13)	7.66×10^{-9}	−0.14	New
	rs79948748	12	57,873,498		T	0.89	1.08 (1.04–1.13)	5.30×10^{-4}	1.10 (1.05–1.15)	2.41×10^{-5}		
12q24.11	rs17550549	12	111,357,471	MYL2/SH2B3	T	0.20	0.91 (0.88–0.95)	1.61×10^{-7}	0.91 (0.88–0.94)	2.81×10^{-8}	−0.08	New
	rs78894077	12	111,856,673		T	0.06	0.86 (0.80–0.92)	2.22×10^{-5}	0.85 (0.79–0.91)	6.15×10^{-6}		
16q23.2	rs140851213	16	79,754,433	WWOX, MAF	T	0.28	0.92 (0.88–0.96)	3.57×10^{-5}	0.91 (0.88–0.95)	7.66×10^{-6}	−0.02	New
	rs4341754	16	80,039,621		C	0.43	0.91 (0.89–0.94)	5.12×10^{-12}	0.91 (0.89–0.93)	2.29×10^{-12}		

Locus	SNP	CHR	Pos	Nearest gene	EA	Single-SNP meta-analysis			Joint analysis			Multiple-independent-variant loci New/ known[b]
						Freq	OR (95% CI)	P	OR (95% CI)	P	r^a	
17p13.3	rs9915645	17	812,534	NXN	T	0.48	0.92 (0.90–0.95)	4.87×10^{-10}	0.93 (0.90–0.95)	2.13×10^{-9}	−0.02	New
	rs11651883	17	835,502		T	0.40	1.06 (1.03–1.09)	1.47×10^{-5}	1.06 (1.03–1.09)	1.25×10^{-5}		
17p12	rs368674461	17	10,485,457	MYHAS/TMEM238L	T	0.01	0.62 (0.49–0.79)	7.85×10^{-5}	0.60 (0.47–0.76)	2.64×10^{-5}	0.04	New
	rs1078643	17	10,707,241		A	0.77	1.13 (1.09–1.17)	2.05×10^{-13}	1.13 (1.09–1.17)	1.59×10^{-13}		
20p12.3	rs6117209	20	6,302,114	BMP2	T	0.80	1.08 (1.04–1.12)	1.84×10^{-5}	1.09 (1.05–1.13)	9.93×10^{-6}	0.00	New
	rs6085662	20	6,698,372		C	0.21	1.10 (1.07–1.14)	1.69×10^{-9}	1.10 (1.07–1.14)	1.06×10^{-9}		

Abbreviations: CHR, chromosome; CI, confidence interval; EA, effective allele; Freq, frequency of the effective allele; Pos, position (hg19); *r*, linkage disequilibrium correlation between an SNP and the next adjacent SNP at a locus.

[a] LD correlation between an SNP and the next adjacent SNP at a locus.

[b] Known: loci known to contain independent secondary signals for CRC risk in East Asians; new: loci not reported before to contain independent secondary signals for CRC risk in East Asians.